(2.2) Definition of pointwise vector operations:

(a) The **sum** $f+g$ of $f,g \in \mathbb{F}^T$ is the function

$$f+g : T \to \mathbb{F} : t \mapsto f(t)+g(t).$$

(s) The **product** αf of the **scalar** $\alpha \in \mathbb{F}$ with the function $f \in \mathbb{F}^T$ is the function

$$\alpha f : T \to \mathbb{F} : t \mapsto \alpha f(t).$$

With respect to these operations, $\mathbb{F}^T$ is a linear space (over $\mathbb{F}$). In particular, the function

$$0 : T \to \mathbb{F} : t \mapsto 0$$

is the neutral element, or zero vector, and, for $f \in \mathbb{F}^T$,

$$-f : T \to \mathbb{F} : t \mapsto -f(t)$$

is the additive inverse for f.

Note that it is not possible to add two functions unless they have the same domain!

Standard examples include:

(i) $T = \underline{n}$, in which case we get n-**dimensional coordinate space** $\mathbb{F}^n$ whose elements (vectors) we call n-vectors.

(ii) $T = \underline{m} \times \underline{n}$, in which case we get the space $\mathbb{F}^{m\times n}$, whose elements we call m-by-n matrices.

(iii) $T = \mathbb{R}$, $\mathbb{F} = \mathbb{R}$, in which case we get the space of all real-valued functions on the real line.

(iv) $T = \mathbb{R}^n$, $\mathbb{F} = \mathbb{R}$, in which case we get the space of all real-valued functions of n real variables.

The most common way to get a vector space is as a *linear subspace*:

Definition: A nonempty subset Y of a vector space X is a **linear subspace** (of X) in case it is closed under addition and multiplication by a scalar. This means that the two sets

$$Y+Y := \{y+z : y,z \in Y\} \quad \text{and} \quad \mathbb{F}Y := \{\alpha y : \alpha \in \mathbb{F}, y \in Y\}$$

are in Y.

Standard examples include:

(i) The **trivial space** $\{0\}$, consisting of the zero vector alone; it's a great space for testing one's understanding.

(ii) $\Pi_k :=$ the set of all **polynomials of degree** $\le k$ as a subset of $\mathbb{F}^{\mathbb{F}}$.

(iii) The set $C[a\,..\,b]$ of all **continuous functions** on the interval $[a\,..\,b]$.

(iv) The set of all real symmetric matrices of order n as a subset of $\mathbb{R}^{n\times n}$.

(v) The set of all real-valued functions on $\mathbb{R}$ that vanish on some fixed set S.

(vi) The set $\mathrm{BL}_\xi \subset C[\xi_1\,..\,\xi_{\ell+1}]$ of all **broken lines** with (interior) breaks at $\xi_2 < \cdots < \xi_\ell$.

It is a good exercise to check that, according to the abstract definition of a vector space, any linear subspace of a vector space is again a vector space. Conversely, if a subset of a vector space is *not* closed under vector addition or under multiplication by scalars, then it cannot be a vector space (with respect to the given operations) since it violates the basic assumption that the sum of any two elements and the product of any scalar with any element is again an element of the space. (To be sure, the empty subset $\{\}$ of a linear space is vacuously closed under the two vector operations but fails to be a linear subspace since it fails to be nonempty.)

Proposition: A subset Y of a vector space X is a vector space (with respect to the same addition and multiplication by scalars) if and only if Y is a linear subspace (of X), i.e., Y is nonempty and is closed under addition and multiplication by scalars.

Corollary: The sum, $Y + Z := \{y + z : y \in Y, z \in Z\}$, and the intersection, $Y \cap Z$, of two linear subspaces, Y and Z, of a vector space is a linear subspace.

We saw that pointwise addition and multiplication by a scalar makes the collection $\mathbb{F}^T$ of all maps from some set T to the scalars a vector space. The same argument shows that the collection X^T of all maps from some set T into a *vector space* X (over the scalar field $\mathbb{F}$) is a vector space under pointwise addition and multiplication by scalars. This means, explicitly, that we define the sum $f + g$ of $f, g \in X^T$ by

$$f + g : T \to X : t \mapsto f(t) + g(t)$$

and define the product αf of $f \in X^T$ with the scalar $\alpha \in \mathbb{F}$ by

$$\alpha f : T \to X : t \mapsto \alpha f(t).$$

Thus, we can generate from one vector space X many different vector spaces, namely all the linear subspaces of the vector space X^T, with T an arbitrary set.

2.1 For each of the following sets of real-valued assignments or maps, determine whether or not they form a vector space (with respect to pointwise addition and multiplication by scalars), and give a reason for your answer. (a) $\{x \in \mathbb{R}^3 : x_1 = 4\}$; (b) $\{x \in \mathbb{R}^3 : x_1 = x_2\}$; (c) $\{x \in \mathbb{R}^3 : 0 \le x_j, j = 1, 2, 3\}$; (d) $\{(0, 0, 0)\}$; (e) $\{x \in \mathbb{R}^3 : x \notin \mathbb{R}^3\}$; (f) $C[0\,..\,2]$; (g) The collection of all 3×3 matrices with all diagonal entries equal to zero.

2.2 Prove that, for every x in the vector space X, $(-1)x = -x$, and $0x = 0$.

2.3 Prove that the intersection of any collection of linear subspaces of a vector space is a linear subspace.

2.4 Prove: The union of two linear subspaces is a linear subspace if and only if one of them contains the other.

2.5 Provide a proof of the above Proposition.

Linear maps

Definition: Let X, Y be vector spaces (over the same scalar field $\mathbb{F}$). The map $f : X \to Y$ is called **linear** if it is

(a) **additive**, i.e.,

$$\forall\{x, z \in X\}\ f(x + z) = f(x) + f(z);$$

and

(s) **homogeneous**, i.e.,

$$\forall\{x \in X, \alpha \in \mathbb{F}\}\ f(\alpha x) = \alpha f(x).$$

We denote the collection of all linear maps from X to Y by

$$L(X, Y).$$

Many books call a linear map a **linear transformation** or a **linear operator**. It is customary to denote linear maps by capital letters. Further, if A is a linear map and $x \in \operatorname{dom} A$, then it is customary to write Ax instead of $A(x)$.

Examples: If X is a linear subspace of $\mathbb{F}^T$, then, for every $t \in T$, the map

$$\delta_t : X \to \mathbb{F} : f \mapsto f(t)$$

of evaluation at t is linear since the vector operations are pointwise.

The map $D : C^{(1)}(\mathbb{R}) \to C(\mathbb{R}) : g \mapsto Dg$ that associates with each continuously differentiable function g its first derivative Dg is a linear map.

The map $C[a \,..\, b] \to \mathbb{R} : g \mapsto \int_a^b g(t)\,\mathrm{d}t$ is linear.

Let $\mathbf{c} := \{a : \mathbb{N} \to \mathbb{F} : \lim_{n\to\infty} a_n \text{ exists}\}$, i.e., $\mathbf{c}$ is the vector space of all convergent sequences. Then the map $\mathbf{c} \to \mathbb{F} : a \mapsto \lim_{n\to\infty} a_n$ is linear.

These examples show that the basic operations in Calculus are linear. This is the reason why so many people outside Algebra, such as Analysts and Applied Mathematicians, are so interested in Linear Algebra.

The simplest linear map on a vector space X to a vector space Y is the so-called **trivial map**. It is the linear map that maps every element of X to 0; it is, itself, denoted by

$$0.$$

It is surprising how often this map serves as a suitable illustration or counterexample.

Example: If $a \in \mathbb{R}^n$, then

$$a^{\mathrm{t}} : \mathbb{R}^n \to \mathbb{R} : x \mapsto a^{\mathrm{t}}x := a_1x_1 + a_2x_2 + \cdots + a_nx_n \tag{2.3}$$

is a linear map of great practical importance. Indeed, any (real) linear algebraic equation in n unknowns has the form

$$a^{\mathrm{t}}? = b$$

for some **coefficient vector** $a \in \mathbb{R}^n$ and some **right side** $b \in \mathbb{R}$. Such an equation has solutions for arbitrary b if and only if $a \neq 0$. You have already learned that the general solution can always be written as the sum of a particular solution and an arbitrary solution of the corresponding **homogeneous** equation

$$a^{\mathrm{t}}? = 0.$$

In particular, the map a^{t} cannot be 1-1 unless $n = 1$.

Assume that $a \neq 0$. For $n = 2$, it is instructive and easy to visualize the solution set as a straight line, parallel to the straight line

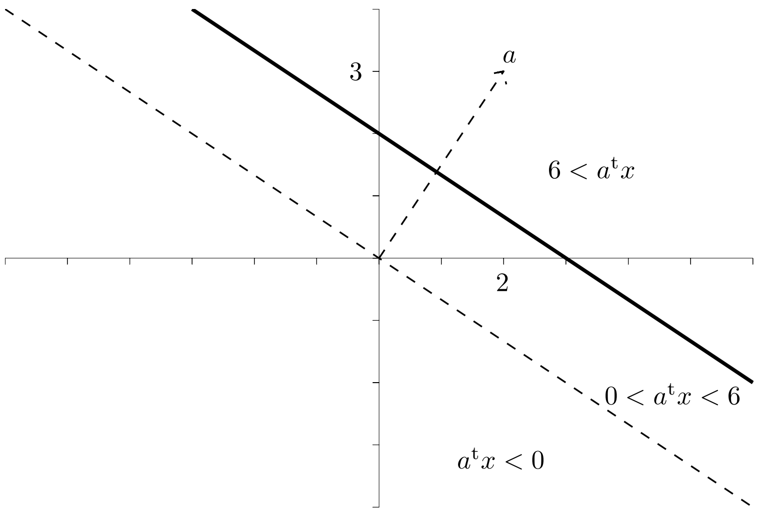
$$\operatorname{null} a^{\mathrm{t}} := \{x \in \mathbb{R}^2 : a^{\mathrm{t}}x = 0\}$$

through the origin formed by all the solutions to the corresponding homogeneous problem, and perpendicular to the coefficient vector a. Note that the 'nullspace' $\operatorname{null} a^{\mathrm{t}}$ splits $\mathbb{R}^2$ into the two **half-spaces**

$$\{x \in \mathbb{R}^2 : a^{\mathrm{t}}x > 0\} \qquad \{x \in \mathbb{R}^2 : a^{\mathrm{t}}x < 0\},$$

one of which contains a. Here is such a figure, for the particular equation

$$2x_1 + 3x_2 = 6.$$



(2.4) Figure. One way to visualize all the parts of the equation $a^t x = 6$ with $a = (2, 3)$.

□

By adding or composing two linear maps (if appropriate) or by multiplying a linear map by a scalar, we obtain further linear maps. Here are the details.

The (pointwise) sum $A+B$ of $A, B \in L(X, Y)$ and the product αA of $\alpha \in \mathbb{F}$ with $A \in L(X, Y)$ are again in $L(X, Y)$, hence $L(X, Y)$ is closed under (pointwise) addition and multiplication by a scalar, therefore a linear subspace of the vector space Y^X of all maps from X into the vector space Y.

$L(X, Y)$ is a vector space under pointwise addition and multiplication by a scalar.

Linearity is preserved not only under (pointwise) addition and multiplication by a scalar, but also under map *composition*.

The composition of two linear maps is again linear (if it is defined).

Indeed, if $A \in L(X, Y)$ and $B \in L(Y, Z)$, then BA maps X to Z and, for any $x, y \in X$,

$$(BA)(x + y) = B(A(x + y)) = B(Ax + Ay) = B(Ax) + B(Ay) = (BA)(x) + (BA)(y).$$

Also, for any $x \in X$ and any scalar α,

$$(BA)(\alpha x) = B(A(\alpha x)) = B(\alpha Ax) = \alpha B(Ax) = \alpha(BA)(x).$$

2.6 For each of the following maps, determine whether or not it is linear (give a reason for your answer).

(a) $\Pi_{<k} \to \mathbb{N} : p \mapsto \#\{x : p(x) = 0\}$ (i.e., the map that associates with each polynomial of degree $< k$ the number of its zeros).

(b) $C[a \,..\, b] \to \mathbb{R} : f \mapsto \max_{a \le x \le b} f(x)$

(c) $\mathbb{F}^{3\times 4} \to \mathbb{F} : A \mapsto A(2, 2)$

(d) $L(X, Y) \to Y : A \mapsto Ax$, with x a fixed element of X (and, of course, X and Y vector spaces).

(e) $\mathbb{R}^{m\times n} \to \mathbb{R}^{n\times m} : A \mapsto A^c$ (with A^c the (conjugate) transpose of the matrix A)

(f) $\mathbb{R} \to \mathbb{R}^2 : x \mapsto (x, \sin(x))$

2.7 *The linear image of a vector space is a vector space*: Let $f : X \to T$ be a map on some vector space X into some set T on which addition and multiplication by scalars is defined in such a way that

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y), \qquad \alpha, \beta \in \mathbb{F},\ x, y \in X. \tag{2.5}$$

Prove that $\operatorname{ran} f$ is a vector space (with respect to the addition and multiplication as restricted to $\operatorname{ran} f$). (See Problem 4.17 for an important application.)

Linear maps from $\mathbb{F}^n$

As a ready source of many examples, we now give a complete description of $L(\mathbb{F}^n, X)$.

For any sequence $v_1, v_2, \ldots, v_n$ in the vector space X, the map

$$f : \mathbb{F}^n \to X : a \mapsto v_1 a_1 + v_2 a_2 + \cdots + v_n a_n$$

is linear.

Proof: The proof is a boring but necessary verification.

(a) additivity:

$$\begin{aligned}
f(a+b) &= v_1\,(a+b)_1 + v_2\,(a+b)_2 + \cdots + v_n\,(a+b)_n && \text{(definition of } f\text{)}\\
&= v_1\,(a_1+b_1) + v_2\,(a_2+b_2) + \cdots + v_n\,(a_n+b_n) && \text{(addition of } n\text{-vectors)}\\
&= v_1a_1 + v_1b_1 \ + \ v_2a_2 + v_2b_2 \ + \ \cdots \ + \ v_na_n + v_nb_n && \text{(multipl. by scalar distributes)}\\
&= v_1a_1 + v_2a_2 + \cdots + v_na_n \quad + \quad v_1b_1 + v_2b_2 + \cdots + v_nb_n && \text{(vector addition commutes)}\\
&= f(a) \ + \ f(b) && \text{(definition of } f\text{)}
\end{aligned}$$

(s) homogeneity:

$$\begin{aligned}
f(\alpha a) &= v_1\,(\alpha a)_1 + v_2\,(\alpha a)_2 + \cdots + v_n\,(\alpha a)_n && \text{(definition of } f\text{)}\\
&= v_1\alpha a_1 + v_2\alpha a_2 + \cdots + v_n\alpha a_n && \text{(multipl. of scalar with } n\text{-vectors)}\\
&= \alpha(v_1a_1 + v_2a_2 + \cdots + v_na_n) && \text{(multipl. by scalar distributes)}\\
&= \alpha f(a) && \text{(definition of } f\text{)}
\end{aligned}$$

□

Definition: The weighted sum

$$v_1a_1 + v_2a_2 + \cdots + v_na_n$$

is called the **linear combination of the vectors** $v_1, v_2, \ldots, v_n$ **with weights** $a_1, a_2, \ldots, a_n$. I will use the suggestive abbreviation

$$[v_1, v_2, \ldots, v_n]a := v_1a_1 + v_2a_2 + \cdots + v_na_n,$$

hence use

$$[v_1, v_2, \ldots, v_n]$$

for the map $\mathbb{F}^n \to X : a \mapsto v_1a_1 + v_2a_2 + \cdots + v_na_n$. I call such a map a **column map**, and call v_j its j**th column**. Further, I denote its number of columns by

$$\#V.$$

The most important special case of this occurs when also X is a coordinate space, $X = \mathbb{F}^m$ say. In this case, each v_j is an m-vector, and

$$v_1 a_1 + v_2 a_2 + \cdots + v_n a_n = V a,$$

with V the $m \times n$-matrix with columns $v_1, v_2, \ldots, v_n$. This explains why I chose to write the weights in the linear combination $v_1 a_1 + v_2 a_2 + \cdots + v_n a_n$ to the right of the vectors v_j rather than to the left. For, it suggests that working with the map $[v_1, v_2, \ldots, v_n]$ is rather like working with a matrix with columns $v_1, v_2, \ldots, v_n$.

Note that `MATLAB` uses the notation $[v_1, v_2, \ldots, v_n]$ for the matrix with columns $v_1, v_2, \ldots, v_n$, as do some textbooks. This stresses the fact that it is customary to think of the *matrix* $C \in \mathbb{F}^{m \times n}$ with columns $c_1, c_2, \ldots, c_n$ as the *linear map* $[c_1, c_2, \ldots, c_n] : \mathbb{F}^n \to \mathbb{F}^m : x \mapsto c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$.

□

Agreement: For any sequence $v_1, v_2, \ldots, v_n$ of m-vectors,

$$[v_1, v_2, \ldots, v_n]$$

denotes both the $m \times n$-matrix V with columns $v_1, v_2, \ldots, v_n$ and the linear map

$$V : \mathbb{F}^n \to \mathbb{F}^m : a \mapsto [v_1, v_2, \ldots, v_n] a = v_1 a_1 + v_2 a_2 + \cdots + v_n a_n.$$

Thus,

$$\mathbb{F}^{m \times n} = L(\mathbb{F}^n, \mathbb{F}^m).$$

Thus, a matrix $V \in \mathbb{F}^{m \times n}$ is associated with two rather different maps: (i) since it is an assignment with domain $\underline{m} \times \underline{n}$ and values in $\mathbb{F}$, we could think of it as a map on $\underline{m} \times \underline{n}$ to $\mathbb{F}$; (ii) since it is the n-list of its columns, we can think of it as the linear map from $\mathbb{F}^n$ to $\mathbb{F}^m$ that carries the n-vector a to the m-vector $Va = v_1 a_1 + v_2 a_2 + \cdots + v_n a_n$. From now on, we will stick to the second interpretation when we talk about the domain, the range, or the target, of a matrix. Thus, for $V \in \mathbb{F}^{m \times n}$, $\operatorname{dom} V = \mathbb{F}^n$ and $\operatorname{tar} V = \mathbb{F}^m$, and $\operatorname{ran} V \subset \mathbb{F}^m$. – If we want the first interpretation, we call $V \in \mathbb{F}^{m \times n}$ a (two-dimensional) **array**.

Next, we prove that there is nothing special about the linear maps of the form $[v_1, v_2, \ldots, v_n]$ from $\mathbb{F}^n$ into the vector space X, i.e., *every* linear map from $\mathbb{F}^n$ to X is necessarily of that form. The identity map

$$\mathrm{id}_n : \mathbb{F}^n \to \mathbb{F}^n : a \to a$$

is of this form, i.e.,

$$\mathrm{id}_n = [e_1, e_2, \ldots, e_n]$$

with e_j the j**th unit vector**, i.e.,

$$e_j := (\underbrace{0, \ldots, 0}_{j-1 \text{ zeros}}, 1, 0, \ldots, 0)$$

the vector (with the appropriate number of entries) all of whose entries are 0, except for the jth, which is 1. Written out in painful detail, this says that

$$a = e_1 a_1 + e_2 a_2 + \cdots + e_n a_n, \qquad \forall a \in \mathbb{F}^n.$$

Further,

(2.6) Proposition: If $V = [v_1, v_2, \ldots, v_n] : \mathbb{F}^n \to X$ and $f \in L(X, Y)$, then $fV = [f(v_1), \ldots, f(v_n)]$.

Proof: If $\operatorname{dom} f = X$ and f is linear, then fV is linear and, for any $a \in \mathbb{F}^n$,

$$(fV)a = f(Va) = f(v_1a_1 + v_2a_2 + \cdots + v_na_n) = f(v_1)a_1 + f(v_2)a_2 + \cdots + f(v_n)a_n = [f(v_1), \ldots, f(v_n)]a.$$

□

Consequently, for any $f \in L(\mathbb{F}^n, X)$,

$$f = f \ \mathrm{id}_n = f\,[e_1, e_2, \ldots, e_n] = [f(e_1), \ldots, f(e_n)].$$

This proves:

(2.7) Proposition: The map f from $\mathbb{F}^n$ to the vector space X is linear if and only if

$$f = [f(e_1), f(e_2), \ldots, f(e_n)].$$

In other words,

$$L(\mathbb{F}^n, X) = \{[v_1, v_2, \ldots, v_n] : v_1, v_2, \ldots, v_n \in X\} \qquad (\simeq X^n).$$

As a simple example, recall from (2.3) the map $a^{\mathrm{t}} : \mathbb{R}^n \to \mathbb{R} : x \mapsto a_1x_1 + a_2x_2 + \cdots + a_nx_n = [a_1, \ldots, a_n]x$, and, in this case, $a^{\mathrm{t}}e_j = a_j$, all j. This confirms that a^{t} is linear and shows that

$$a^{\mathrm{t}} = [a_1, \ldots, a_n] = [a]^{\mathrm{t}}. \tag{2.8}$$

Notation: I follow `MATLAB` notation. E.g., $[V, W]$ denotes the column map in which first all the columns of V are used and then all the columns of W. Also, if V and W are column maps, then I write

$$V \subset W$$

to mean that V is obtained by omitting (zero or more) columns from W; i.e., $V = W(:, c)$ for some subsequence c of $1{:}\#W$.

Finally, if W is a column map and M is a set, then I'll write

$$W \ \subset \ M$$

to mean that the columns of W are elements of M. For example:

(2.9) Proposition: If Z is a linear subspace of Y and $W \in L(\mathbb{F}^m, Y)$, then $W \subset Z \implies \operatorname{ran} W \subset Z$.

The important (2.6)Proposition is the reason we define the **product of matrices** the way we do, namely as

$$(AB)(i,j) := \sum_k A(i,k)B(k,j), \quad \forall i, j.$$

For, if $A \in \mathbb{F}^{m\times n} = L(\mathbb{F}^n, \mathbb{F}^m)$ and $B = [b_1, b_2, \ldots, b_r] \in \mathbb{F}^{n\times r} = L(\mathbb{F}^r, \mathbb{F}^n)$, then $AB \in L(\mathbb{F}^r, \mathbb{F}^m) = \mathbb{F}^{m\times r}$, and

$$AB = A[b_1, b_2, \ldots, b_r] = [Ab_1, \ldots, Ab_r].$$

Notice that the product AB of two maps A and B makes sense if and only if $\operatorname{dom} A \supset \operatorname{tar} B$. For matrices A and B, this means that the number of columns of A must equal the number of rows of B; we couldn't apply A to the columns of B otherwise.

In particular, *the 1-column matrix* $[Ax]$ *is the product of the matrix* A *with the 1-column matrix* $[x]$, i.e.,

$$A[x] = [Ax], \quad \forall A \in \mathbb{F}^{m\times n}, x \in \mathbb{F}^n.$$

For this reason, most books on elementary linear algebra and most users of linear algebra *identify* the n-vector x with the $n \times 1$-matrix $[x]$, hence write simply x for what I have denoted here by $[x]$. I will feel free from now on to use the same identification. However, I will not be doctrinaire about it. In particular, I will continue to specify a particular n-vector x by writing down its entries in a list, like $x = (x_1, x_2, \ldots)$, since that uses much less space than does the writing of

$$[x] = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}.$$

It is consistent with the standard identification of the n-vector x with the $n \times 1$-matrix $[x]$ to mean by x^{t} the $1 \times n$-matrix $[x]^{\mathrm{t}}$. Further, with y also an n-vector, one identifies the $(1,1)$-*matrix* $[x]^{\mathrm{t}}[y] = x^{\mathrm{t}}y$ with the *scalar*

$$\sum_j x_j y_j = y^{\mathrm{t}}x.$$

On the other hand,

$$yx^{\mathrm{t}} = [y][x]^{\mathrm{t}} \ = \ (y_i x_j : (i,j) \in \underline{n} \times \underline{n})$$

is an $n \times n$-*matrix* (and identified with a scalar only if $n = 1$).

However, I will *not* use the terms 'column vector' or 'row vector', as they don't make sense to me. Also, whenever I want to stress the fact that x or x^{t} is meant to be a matrix, I will write $[x]$ and $[x]^{\mathrm{t}}$, respectively.

For example, what about the expression $xy^{\mathrm{t}}z$ in case x, y, and z are vectors? It makes sense only if y and z are vectors of the same length, say $y, z \in \mathbb{F}^n$. In that case, it is $[x][y]^{\mathrm{t}}[z]$, and this we can compute in two ways: we can apply the matrix xy^{t} to the vector z, or we can multiply the vector x with the scalar $y^{\mathrm{t}}z$. Either way, we obtain the vector $x(y^{\mathrm{t}}z) = (y^{\mathrm{t}}z)x$, i.e., the $(y^{\mathrm{t}}z)$-multiple of x. However, while the product $x(y^{\mathrm{t}}z)$ of x with $(y^{\mathrm{t}}z)$ makes sense both as a matrix product and as multiplication of the vector x by the scalar $y^{\mathrm{t}}z$, the product $(y^{\mathrm{t}}z)x$ *only* makes sense as a product of the scalar $y^{\mathrm{t}}z$ with the vector x.

(2.10) Example: Here is an example, of help later. Consider the socalled **elementary row operation**

$$E_{i,k}(\alpha)$$

on n-vectors, in which one adds α times the kth entry to the ith entry. Is this a linear map? What is a formula for it?

We note that the kth entry of any n-vector x can be computed as $e_k{}^{\mathrm{t}}x$, while adding β to the ith entry of x is accomplished by adding βe_i to x. Hence, adding α times the kth entry of x to its ith entry replaces x by $x + e_i(\alpha e_k{}^{\mathrm{t}}x) = x + \alpha e_i e_k{}^{\mathrm{t}}x$. This gives the handy formula

$$E_{i,k}(\alpha) = \ \mathrm{id}_n + \alpha e_i e_k{}^{\mathrm{t}}. \tag{2.11}$$

Now, to check that $E_{i,j}(\alpha)$ is linear, we observe that it is the sum of two maps, and the first one, id_n, is certainly linear, while the second is the composition of the three maps,

$$e_k{}^{\mathrm{t}} : \mathbb{F}^n \to \mathbb{F} \simeq \mathbb{F}^1 : z \mapsto e_k{}^{\mathrm{t}}z, \quad [e_i] : \mathbb{F}^1 \to \mathbb{F}^n : \beta \to e_i\beta, \quad \alpha : \mathbb{F}^n \to \mathbb{F}^n : z \mapsto \alpha z,$$

and each of these is linear (the last one because we assume $\mathbb{F}$ to be a *commutative* field).

Matrices of the form

$$E_{y,z}(\alpha) := \ \mathrm{id} + \alpha y z^{\mathrm{t}} \tag{2.12}$$

are called **elementary**. They are very useful since, if invertible, their inverse has the same simple form; see (2.19)Proposition below. □

2.8 Use the fact that the jth column of the matrix A is the image of e_j under the linear map A to construct the matrices that carry out the given action.

(i) The matrix A of order 2 that rotates the plane clockwise 90 degrees;

(ii) The matrix B that reflects $\mathbb{R}^n$ across the hyperplane $\{x \in \mathbb{R}^n : x_n = 0\}$;

(iii) The matrix C that keeps the hyperplane $\{x \in \mathbb{R}^n : x_n = 0\}$ pointwise fixed, and maps e_n to $-e_n$;

(iv) The matrix D of order 2 that keeps the y-axis fixed and maps $(1,1)$ to $(2,1)$.

2.9 Use the fact that the jth column of the matrix $A \in \mathbb{F}^{m\times n}$ is the image of e_j under A to derive the four matrices A^2, AB, BA, and B^2 for each of the given pair A and B: (i) $A = [e_1, 0]$, $B = [0, e_1]$; (ii) $A = [e_2, e_1]$, $B = [e_2, -e_1]$; (iii) $A = [e_2, e_3, e_1]$, $B = A^2$.

2.10 For each of the following pairs of matrices A, B, determine their products AB and BA if possible, or else state why it cannot be done.

(a) $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, B the matrix `eye(2)`; (b) $A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 2 \end{bmatrix}$, $B = A^{\mathrm{t}}$; (c) $A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$;
(d) $A = \begin{bmatrix} 2+\mathrm{i} & 4-\mathrm{i} \\ 3-\mathrm{i} & 3+\mathrm{i} \end{bmatrix}$, $B = \begin{bmatrix} 2-\mathrm{i} & 3+\mathrm{i} & 3\mathrm{i} \\ 3-\mathrm{i} & 4+\mathrm{i} & 2 \end{bmatrix}$.

2.11 For any $A, B \in L(X)$, the products AB and BA are also linear maps on X, as are $A^2 := AA$ and $B^2 := BB$. Give an example of $A, B \in L(X)$ for which $(A+B)^2$ does not equal $A^2 + 2AB + B^2$. (Hint: keep it as simple as possible, by choosing X to be $\mathbb{R}^2$, hence both A and B are 2-by-2 matrices.)

2.12 Give an example of matrices A and B for which both $AB = 0$ and $BA = 0$, while neither A nor B is a zero matrix.

2.13 Prove: *If A and B are matrices with the same number of rows, and C and D are such that AC and BD are defined, then the product of the two partitioned matrices $[A, B]$ and $[C; D]$ is defined and equals $AB + CD$.*

2.14 Prove that both $\mathbb{C} \to \mathbb{R} : z \mapsto \operatorname{Re} z$ and $\mathbb{C} \to \mathbb{R} : z \mapsto \operatorname{Im} z$ are linear maps when we consider $\mathbb{C}$ as a vector space over the real scalar field.

The linear equation $A? = y$, and $\operatorname{ran} A$ and $\operatorname{null} A$

We are ready to recognize and use the fact that the general system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ \cdots &= \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \end{aligned} \tag{2.13}$$

of m linear equations in the n unknowns $x_1, \ldots, x_n$ is equivalent to the vector equation

$$Ax = y,$$

provided

$$x := (x_1, \ldots, x_n), \quad y := (y_1, \ldots, y_m), \quad A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Here, **equivalence** means that the entries $x_1, \ldots, x_n$ of the n-vector x solve the system of linear equations (2.13) if and only if x solves the vector equation $A? = y$. This equivalence is not only a notational convenience. Switching from (2.13) to $A? = y$ is the conceptual shift that started Linear Algebra. It shifts the focus, from the scalars $x_1, \ldots, x_n$, to the vector x formed by them, and to the map A given by the coefficients in (2.13), its range and nullspace (about to be defined), and this makes for simplicity, clarity, and generality.

To stress the generality, we now give a preliminary discussion of the equation

$$A? = y$$

in case A is a *linear* map, from the vector space X to the vector space Y say, with y some element of Y.

Existence of a solution for every $y \in Y$ is equivalent to having A be *onto*, i.e., to having $\operatorname{ran} A = Y$. Now, the range of A is the linear image of a vector space, hence itself a vector space. Indeed, if $v_1, v_2, \ldots, v_m$

are elements of $\operatorname{ran} A$, then there must be a sequence $w_1, w_2, \ldots, w_m$ in X with $Aw_j = v_j$, all j. Since X is a vector space, it contains Wa for arbitrary $a \in \mathbb{F}^m$, therefore the corresponding linear combination $Va = [Aw_1, Aw_2, \ldots, Aw_m]a = (AW)a = A(Wa)$ must be in $\operatorname{ran} A$. In other words, if $V \subset \operatorname{ran} A$, then $\operatorname{ran} V \subset \operatorname{ran} A$.

Hence, if we wonder whether A is onto, and we happen to know an *onto* column map $[v_1, v_2, \ldots, v_m] = V \in L(\mathbb{F}^m, Y)$, then we only have to check that the *finitely many* columns, $v_1, v_2, \ldots, v_m$, of V are in $\operatorname{ran} A$. For, if some are not in $\operatorname{ran} A$, then, surely, A is not onto. However, if they all are in $\operatorname{ran} A$, then $Y = \operatorname{ran} V \subset \operatorname{ran} A \subset \operatorname{tar} A = Y$, hence $\operatorname{ran} A = Y$ and A is onto.

(2.14)Proposition: The range of a linear map $A \in L(X, Y)$ is a linear subspace, i.e., is nonempty and closed under vector addition and multiplication by a scalar.

If Y is the range of the column map V, then A is onto if and only if the finitely many columns of V are in $\operatorname{ran} A$.

Uniqueness of a solution for every $y \in Y$ is equivalent to having A be *1-1*, i.e., to have $Ax = Az$ imply that $x = z$. For a *linear* map $A : X \to Y$, we have $Ax = Az$ if and only if $A(x - z) = 0$. In other words, if $y = Ax$, then

$$A^{-1}\{y\} = x + \{z \in X : Az = 0\}. \tag{2.15}$$

In particular, A is 1-1 if and only if $\{z \in X : Az = 0\} = \{0\}$. In other words, to check whether a *linear* map is 1-1, we only have to check whether it is 1-1 'at' one particular point, e.g., 'at' 0. For this reason, the set $\{z \in X : Az = 0\}$ of all elements of X mapped by A to 0 is singled out.

Definition: The set

$$\operatorname{null} A := \{z \in \operatorname{dom} A : Az = 0\}$$

is called the **nullspace** or **kernel** of the linear map A.
The linear map is 1-1 if and only if its nullspace is **trivial**, i.e., contains only the zero vector.
The nullspace of a linear map is a linear subspace.

Almost all linear subspaces you'll meet will be of the form $\operatorname{ran} A$ or $\operatorname{null} A$ for some linear map A. These two ways of specifying a linear subspace are very different in character.

If we are told that our linear subspace Z of X is of the form $\operatorname{null} A$, for a certain linear map A on X, then we know, offhand, exactly one element of Z for sure, namely the element 0 which lies in every linear subspace. On the other hand, it is easy to *test* whether a given $x \in X$ lies in $Z = \operatorname{null} A$: simply compute Ax and check whether it is the zero vector.

If we are told that our linear subspace Z of X is of the form $\operatorname{ran} A$ for some linear map from some U into X, then we can 'write down' explicitly every element of $\operatorname{ran} A$: they are all of the form Au for some $u \in \operatorname{dom} A$. On the other hand, it is much harder to *test* whether a given $x \in X$ lies in $Z = \operatorname{ran} A$: Now we have to check whether the equation $A? = x$ has a solution (in U).

As a simple example, the vector space Π_k of all polynomials of degree $\le k$ is usually specified as the range of the column map

$$[()^0, ()^1, \ldots, ()^k] : \mathbb{R}^{k+1} \to \mathbb{R}^{\mathbb{R}},$$

with

$$()^j : \mathbb{R} \to \mathbb{R} : t \mapsto t^j$$

a convenient (though non-standard!) notation for the **monomial of degree** j, i.e., as the collection of all real-valued functions that are of the form

$$t \mapsto a_0 + a_1 t + \cdots + a_k t^k$$

